Physica A 392 (2013) 4154–4159

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Ranking the spreading influence in complex networks

Jian-Guo Liu *, Zhuo-Ming Ren, Qiang Guo

Research Center of Complex Systems Science, University of Shanghai for Science and Technology, Shanghai 200093, PR China

HIGHLIGHTS

- K-shell decomposition method could only identify the network-core successfully.
- The spreading influence of the nodes with same k-shell values are very different.
- We present a ranking way to identify the spreading influence of the nodes with same k-shell values.
- The proposed method could rank the node spreading influence more accurately than the k-shell, degree k and other methods.

ARTICLE INFO

Article history:
Received 25 January 2013
Received in revised form 30 March 2013
Available online 2 May 2013

Keywords:
Network science
Spreading influence
K-shell decomposition

ABSTRACT

Identifying the node spreading influence in networks is an important task to optimally use the network structure and ensure the more efficient spreading in information. In this paper, by taking into account the shortest distance between a target node and the node set with the highest k-core value, we present an improved method to generate the ranking list to evaluate the node spreading influence. Comparing with the epidemic process results for four real networks and the Barabási–Albert network, the parameterless method could identify the node spreading influence more accurately than the ones generated by the degree k, closeness centrality, k-shell and mixed degree decomposition methods. This work would be helpful for deeply understanding the node importance of a network.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Spreading is an ubiquitous process in nature, which describes many important activities in society [1–4], such as the virus spreading [5,6], reaction diffusion processes [7,8], pandemics [9], cascading failures [10] and so on. The knowledge of the spreading pathways through the network of interactions is crucial for developing effective methods to either hinder the disease spreading, or accelerate the information dissemination spreading. So far, a lot of works focus on identifying the most influential spreaders in a network [11–13], for example, the most connected nodes (hubs) are supposed to be the key spreaders, being responsible for the largest scale of the spreading process [14–16]. Recently, Kitsak et al. [1] argued that the node spreading influence is determined by its location in a network. By decomposing a network with the k-shell decomposition method, they found that the most influential nodes, namely the network core, could be identified by the largest k-core values. It should be noticed that the k-shell method assigns many nodes with the same k-core value even though they perform entirely differently in the spreading process. Fig. 1 shows that, for some real networks, there are lots of nodes whose k-core values, denoted as k_s, are equal. By taking into the number of removed and existed links in the decomposition process, Zeng et al. [17] proposed an improved method, named mixed degree decomposition (MDD) method, to distinguish the node spreading influence within the node set with the same k_s value. To different networks, the optimal parameters of the MDD method are determined by the statistical properties of the networks, which hinder its application. By investigating the effects of privileged spreaders on social networks, Borge et al. [12,13] found that the node spreading

* Corresponding author. Tel.: +86 15202121269.
E-mail address: liujg004@ustc.edu.cn (J.-G. Liu).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.04.037

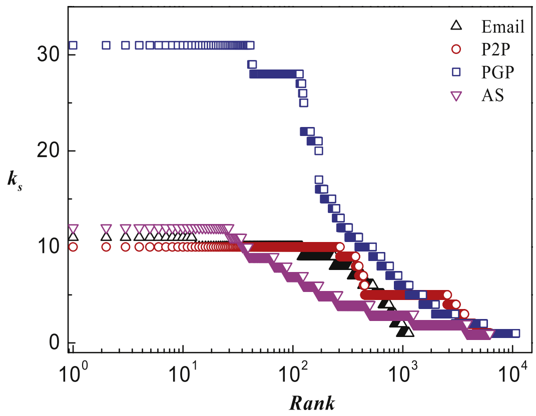

Fig. 1. (Color online) Rank of the k_s values for Email, P2P, PGP, and AS networks, from which one can find that there are lots of nodes whose k_s values are equal.

influence does not depend on their k-core values, which instead determines whether or not a given node prevents the diffusion process. These literatures suggest that, besides the network core, it is also very important to generate a ranking list to identify all nodes' spreading influences. In this paper, we argue that, for the node set with same k-core values, the nodes whose locations are close to the network core have larger spreading influences. Inspired by the idea, we present an improved k-shell method to generate the global influential ranking list. Comparing with the susceptible-infection-recovered (SIR) spreading process [18,19] for four real networks and the Barabási–Albert network [20], the experimental results show that our method could generate the ranking list more accurately than the ones generated by the degree k, closeness centrality(CC) [21], k-shell and MDD decomposition methods respectively.

2. Method

Normally, a network $G = (N, E)$ with N nodes and E links could be described by an adjacent matrix $A = \{a_{ij}\} \in R^{n,n}$, where $a_{ij} = 1$ if node i is connected by node j, and $a_{ij} = 0$ otherwise. The node degree k_i is defined as the number of neighbors for node i. The closeness centrality(CC) of node i is defined as the reciprocal of the sum of the shortest distances to all other nodes of N [21]. The k-shell decomposition method [22,23] could be implemented in the following way to identify the network core. Firstly, remove all nodes with degree one, and then keep pruning the existed nodes until all nodes' degrees are larger than one. The removed nodes would form a node set whose k-core value equals to one. Then, repeat the pruning process in the same way for the rest nodes. Finally, the k-shell method decomposes a network into different node set with different k-core values. Implementing the SIR spreading process for one network, one can find that the nodes with the same k_s values always have different number of infected nodes, namely *spreading influence*. This phenomena suggests that the k-shell decomposition method is not appropriate for ranking the global spreading influence of a network. In terms of the distance from a target node to the network core, the spreading influences of the nodes with the same k-core values could be distinguished in the following way

$$\theta(i|k_s) = (k_s^{\max} - k_s + 1) \sum_{j \in J} d_{ij}, \quad i \in S_{k_s} \tag{1}$$

where $k_s^{\max}$ is the largest k-core value of a network, the shortest distance d_{ij} is measured by the shortest distance from the node i to the node j, J is denoted as the network core node set, and S_{k_s} is denoted as the node set whose k-core values equal to k_s.

3. How to evaluate the performance

To check the performance of the improved method, four real networks are introduced in this paper, which include (i) *Email network* [24]. The Email network of University Rovira i Virgili (URV) of Spain contains faculty, researchers, technicians, managers, administrators, and graduate students. (ii) *Peer-To-Peer (P2P) network* [25]. A sequence of snapshots of the Gnutella peer-to-peer file sharing network from August 2002. Each node represents a host in the Gnutella network and each link represents the connection between each pair of Gnutella hosts. (iii) *Pretty-Good-Privacy (PGP) network* [26]. Pretty-Good-Privacy algorithms have been developed in order to maintain privacy between peers, wherefore, it is also called web of trust of PGP. (iv) *Autonomous Systems (AS) network* [27]. The data was collected from University of Oregon Route Views Project — Online data and reports. The network of routers comprising the Internet can be organized into sub-graphs called Autonomous Systems. Each AS exchanges traffic flows with some neighbors. The statistical properties include the number of nodes N and links E of the network, the average degree, the second-order average degree, and the spreading threshold are given in Table 1.

Download English Version:

https://daneshyari.com/en/article/10481162

Download Persian Version:

https://daneshyari.com/article/10481162

Daneshyari.com